Math 6010
Partial solutions to homework 5

1. We have a linear model with $p = 2$, parameter vector $\boldsymbol{\beta} = (\beta_1 , \beta_2)'$ and design matrix,

$$\boldsymbol{X} = \begin{pmatrix} a_1 & b_1 \\ \vdots & \vdots \\ a_n & b_n \end{pmatrix}.$$

Note that

$$\boldsymbol{X}'\boldsymbol{X} = \begin{pmatrix} \sum_{i=1}^n a_i^2 & \sum_{i=1}^n a_i b_i \\ \sum_{i=1}^n a_i b_i & \sum_{i=1}^n b_i^2 \end{pmatrix},$$

so that

$$(\boldsymbol{X}'\boldsymbol{X})^{-1} = \frac{1}{\left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right) - \left(\sum_{i=1}^n a_i b_i\right)^2} \begin{pmatrix} \sum_{i=1}^n b_i^2 & -\sum_{i=1}^n a_i b_i \\ -\sum_{i=1}^n a_i b_i & \sum_{i=1}^n a_i^2 \end{pmatrix},$$

provided that $\boldsymbol{X}$ has full rank.[1] Since $\hat{\boldsymbol{\beta}} \sim \mathrm{N}_2\left(\boldsymbol{\beta} , \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}\right)$, we can read off the following:

$$\mathrm{Cov}\left(\hat{\beta}_1 , \hat{\beta}_2\right) = -\frac{\sigma^2 \sum_{i=1}^n a_i b_i}{\left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right) - \left(\sum_{i=1}^n a_i b_i\right)^2}.$$

This is zero if and only if $\sum_{i=1}^n a_i b_i = 0$. Because $\hat{\beta}_1$ and $\hat{\beta}_2$ are jointly normal, they are independent if and only if they are uncorrelated; that is, if and only if $\sum_{i=1}^n a_i b_i = 0$.

5. Since $\mathrm{Var}(\hat{\boldsymbol{Y}}) = \boldsymbol{X}\,\mathrm{Var}(\hat{\boldsymbol{\beta}})\boldsymbol{X}' = \sigma^2\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$, the variance of $\hat{Y}_i$ is the (i , i)th element of the preceding variance/covariance matrix; that is,

$$\mathrm{Var}(\hat{Y}_i) = \sigma^2 \sum_{j=1}^p \sum_{k=1}^p X_{i,j} \left[(\boldsymbol{X}\boldsymbol{X})^{-1}\right]_{j,k} X_{i,k}.$$

Because $\sum_{i=1}^n X_{i,j}X_{i,k} = [\boldsymbol{X}'\boldsymbol{X}]_{j,k}$, it follows that

$$\sum_{i=1}^n \mathrm{Var}(\hat{Y}_i) = \sigma^2 \sum_{j=1}^p \sum_{k=1}^p \left[(\boldsymbol{X}\boldsymbol{X})^{-1}\right]_{j,k} [\boldsymbol{X}'\boldsymbol{X}]_{j,k} = \sigma^2 \mathrm{tr}\left((\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{X})\right).$$

Because $(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{X}'\boldsymbol{X}) = \boldsymbol{I}_{p\times p}$, its trace is p; this does the job.

[1] In this case, this means that $(\sum_{i=1}^n a_i^2)(\sum_{i=1}^n b_i^2) \neq (\sum_{i=1}^n a_i b_i)^2$. According to the Cauchy–Schwarz inequality, $(\sum_{i=1}^n a_i^2)(\sum_{i=1}^n b_i^2) \geq (\sum_{i=1}^n a_i b_i)^2$. Therefore, $\boldsymbol{X}$ has full rank if and only if the Cauchy–Schwarz inequality is a strict inequality. This turns out to mean that $z_i := a_i b_i$ is not a linear function of i.

10. We need to assume (as is done in the body of the text) additionally that $\boldsymbol{V}$ is positive definite. In that case, $\boldsymbol{a}'\boldsymbol{V}^{-1}\boldsymbol{a} \geq 0$ for any vector $\boldsymbol{a}$. In particular,

$$F(\boldsymbol{\beta}) := (\boldsymbol{Y} - \boldsymbol{X\beta})'\boldsymbol{V}^{-1}(\boldsymbol{Y} - \boldsymbol{X\beta}) \geq 0.$$

Next, we expand to see that

$$F(\boldsymbol{\beta}) = \boldsymbol{Y}'\boldsymbol{V}^{-1}\boldsymbol{Y} - \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{Y} - \boldsymbol{Y}'\boldsymbol{V}^{-1}\boldsymbol{X\beta} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\beta}.$$

The first and the fourth terms are ≥ 0. The second and the third term are equal to one another because they are scalar quantities (and hence equal to their own transpose). Thus, we see that

$$F(\boldsymbol{\beta}) = \boldsymbol{Y}'\boldsymbol{V}^{-1}\boldsymbol{Y} - 2\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{Y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\beta}.$$

Suppose $\boldsymbol{\beta}_*$ is any solution to $\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\beta}_* = \boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{Y}$. (If $\boldsymbol{X}$ were full rank, then there would be a unique such $\boldsymbol{\beta}_*$. Else, there will be infinitely-many different choices of $\boldsymbol{\beta}_*$; fix any one of them.) By default, $F(\boldsymbol{\beta}_*) = \boldsymbol{Y}'\boldsymbol{V}^{-1}\boldsymbol{Y} - \boldsymbol{\beta}_*'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\beta}_*$. Therefore,

$$\begin{aligned} F(\boldsymbol{\beta}) - F(\boldsymbol{\beta}_*) &= -2\boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{Y} + \boldsymbol{\beta}'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\beta} + \boldsymbol{\beta}_*'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X\beta}_* \\ &= (\boldsymbol{\beta}_* - \boldsymbol{\beta})'\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}(\boldsymbol{\beta}_* - \boldsymbol{\beta}) \\ &\geq 0, \end{aligned}$$

again because $\boldsymbol{V}^{-1}$ is positive definite. This proves that $F(\boldsymbol{\beta})$ is minimized at $\boldsymbol{\beta} = \boldsymbol{\beta}_*$.

15. This problem is an example of a linear model, where the design matrix is

$$\boldsymbol{X} = \begin{pmatrix} 1 & C_1 & S_1 \\ 1 & C_2 & S_2 \\ 1 & C_3 & S_3 \end{pmatrix},$$

where

$$C_i := \cos\left(\frac{2\pi k_1 i}{n}\right), \qquad S_i := \sin\left(\frac{2\pi k_2 i}{n}\right) \qquad \text{for } 1 \leq i \leq n.$$

The theoretical answer is $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$, provided that $\boldsymbol{X}$ has full rank. It is a deep fact from Fourier series that, because $k_1, k_2 \geq 1$ are integers, the columns of $\boldsymbol{X}$ are linearly independent. Next we compute $\widehat{\boldsymbol{\beta}}$.

It might be better to instead proceed directly to find $\widehat{\boldsymbol{\beta}}$ in this case, since the problem is pretty modest in size. Incidentally, the answer in the back of your text is manifestly false, starting with the claim that $\boldsymbol{X}'\boldsymbol{X}$ is a diagonal matrix! In fact, $\boldsymbol{X}'\boldsymbol{X}$ is diagonal only for very special choices of k_1 and k_2. Also, $\widehat{\beta}_0 \neq \overline{Y}$ etc., as we will see soon.

The direct approach to the least squares problem is to solve the minimization problem,

$$\min_{\beta_0, \beta_1, \beta_2} \sum_{i=1}^n [Y_i - \beta_0 - \beta_1 C_i - \beta_2 S_i]^2 := \min_{\beta_0, \beta_1, \beta_2} f(\beta_0\,, \beta_1\,, \beta_2).$$

It is easier to solve this directly, rather than to invert $\boldsymbol{X}'\boldsymbol{X}$, etc. The function f is a positive quadratic, therefore it suffices to set $\partial f/\partial\beta_i = 0$ for $i = 0, 1, 2$. Now,

$$\frac{\partial f}{\partial \beta_0} = -2\sum_{i=1}^{n}\left[Y_i - \beta_0 - \beta_1 C_i - \beta_2 S_i\right] = -2n\left[\overline{Y} - \beta_0 - \beta_1\overline{C} - \beta_2\overline{S}\right],$$

where $\overline{Y} := n^{-1}\sum_{i=1}^{n} Y_i$, $\overline{C} := n^{-1}\sum_{i=1}^{n} C_i$, and $\overline{S} := n^{-1}\sum_{i=1}^{n} S_i$.
Next, we compute

$$\frac{\partial f}{\partial \beta_1} = -2\sum_{i=1}^{n}\left[Y_i - \beta_0 - \beta_1 C_i - \beta_2 S_i\right]C_i = -2n\left[\overline{YC} - \beta_0\overline{C} - \beta_1\overline{C^2} - \beta_2\overline{SC}\right],$$

where $\overline{YC} := n^{-1}\sum_{i=1}^{n} Y_i C_i$, $\overline{C^2} := n^{-1}\sum_{i=1}^{n} C_i^2$ and $\overline{SC} := n^{-1}\sum_{i=1}^{n} S_i C_i$.
Finally,

$$\frac{\partial f}{\partial \beta_2} = -2\sum_{i=1}^{n}\left[Y_i - \beta_0 - \beta_1 C_i - \beta_2 S_i\right]S_i = -2n\left[\overline{YS} - \beta_0\overline{S} - \beta_1\overline{SC} - \beta_2\overline{S^2}\right],$$

where $\overline{YS} := n^{-1}\sum_{i=1}^{n} Y_i S_i$ etc.
Set the preceding three $\partial f/\partial\beta_i$'s equal to zero and solve to obtain a 3×3 linear system,

$$\begin{pmatrix} 1 & \overline{C} & \overline{S} \\ \overline{C} & \overline{C^2} & \overline{SC} \\ \overline{S} & \overline{SC} & \overline{S^2} \end{pmatrix}\widehat{\boldsymbol{\beta}} = \begin{pmatrix} \overline{Y} \\ \overline{YC} \\ \overline{YS} \end{pmatrix}.$$

Therefore,

$$\begin{aligned}\widehat{\boldsymbol{\beta}} &= \begin{pmatrix} 1 & \overline{C} & \overline{S} \\ \overline{C} & \overline{C^2} & \overline{SC} \\ \overline{S} & \overline{SC} & \overline{S^2} \end{pmatrix}^{-1}\begin{pmatrix} \overline{Y} \\ \overline{YC} \\ \overline{YS} \end{pmatrix} \\ &= \frac{1}{\Delta}\begin{pmatrix} \overline{C^2}\cdot\overline{S^2} - (\overline{SC})^2 & \overline{C}\cdot\overline{S^2} - \overline{S}\cdot\overline{SC} & \overline{C}\cdot\overline{SC} - \overline{C^2}\cdot\overline{S} \\ \overline{SC} - \overline{C}\cdot\overline{S} & \overline{S^2} - (\overline{S})^2 & \overline{SC} - \overline{C}\cdot\overline{S} \\ \overline{C}\cdot\overline{SC} - \overline{C^2}\cdot\overline{S} & \overline{S^2} - (\overline{S})^2 & \overline{C^2} - (\overline{C})^2 \end{pmatrix}\begin{pmatrix} \overline{Y} \\ \overline{YC} \\ \overline{YS} \end{pmatrix},\end{aligned}$$

where

$$\Delta = \overline{C^2}\cdot\overline{S^2} - (\overline{SC})^2 - \overline{C}\left(\overline{C}\cdot\overline{S^2} - \overline{S}\cdot\overline{SC}\right) + \overline{S}\left(\overline{C}\cdot\overline{SC} - \overline{S}\cdot\overline{C^2}\right),$$

which is nonzero as we argued, using Fourier series, by other means.